DOCUMENT RESUME

ED 191 901 TM 800 525

AUTHOR Marsh, Herbert W.; Cooper, Terri L.
TITLE Prior Subject Interest, Students' Evaluations, and Instructional Effectiveness.
PUB DATE Apr 80
NOTE 37p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College Faculty; Course Content; *Course Evaluation; Higher Education; Interest Research; Predictor Variables; *Self Evaluation (Individuals); *Student Evaluation of Teacher Performance; *Student Interests; *Teacher Effectiveness

ABSTRACT

The impact of students' prior interest in course subject matter on their evaluations of instructional effectiveness was determined. Students' evaluations of instructional effectiveness were collected in two different university settings involving 1,102 classes, while instructor self evaluations were also collected in one of these settings involving 329 classes. In both university settings, students' prior subject interest was correlated with student ratings of teaching effectiveness, particularly ratings of learning/value and overall course rating. Using data from one of these settings, the combined impact of a wide variety of student, course, and instructor characteristics (expected grade, class size, years in school, teacher rank, etc.) was determined: students' prior subject interest was found to be the most important in predicting their evaluations. In addition, instructors' self evaluations were positively correlated with both their own and their students' perceptions of prior subject interest. Furthermore, the instructors' ratings of learning/value and the overall course were similar to those found with student ratings of teaching effectiveness. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Cooper
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Prior Subject Interest, Students' Evaluations, and Instructional Effectiveness

HERBERT W. MARSH & TERRI L. COOPER

ABSTRACT

The purpose of this paper is to determine the impact of students' prior interest in course subject matter on their evaluations of instructional effectiveness. Data from several previous studies that included a measure of Prior Subject Interest have been analyzed. Students' evaluations of instructional effectiveness were collected in two different university settings (N=1102 classes), while instructor self evaluations of their own teaching were also collected in one of these settings (N=329 classes). In both of the university settings, students' Prior Subject Interest was correlated with student ratings of teaching effectiveness -- particularly ratings of Learning/Value and Overall Course Rating.

Using data from one of these settings the combined impact of a wide variety of student, course, and instructor characteristics (e.g., Expected Grade, Class Size, Year in School, Teacher Rank, etc.) was determined, and students' Prior Subject Interest was was found to be the most important in predicting student evaluations. Furthermore, much of the effect of other background variables could be explained in terms of Prior Subject Interest. For example, a causal model (utilizing path analysis) indicated that approximately 33% of the relationship between Expected Grades and student ratings was eliminated by controlling for Prior Subject Interest.

Instructors in one of the settings were also asked to evaluate their own teaching with the same evaluation form used by students, and to rate their perceptions of their students' Prior Subject Interest. Instructors' self evaluations of their own teaching were positively correlated with both their own and their students' perceptions of Prior Subject Interest. Futhermore, the instructor self evaluations showing the highest correlations with Prior Subject Interest, particularly ratings of Learning/Value and the Overall Course, were similar to those found with student ratings of teaching effectiveness. Since these two distinct measures of teaching effectiveness were similarly related to students' Prior Subject Interest, it appears Prior Interest Subject actually affects the instructional effectiveness in a way that is accurately reflected in the student ratings.

Running Head: Prior Subject Interest

THIS PAPER WAS THE BASIS OF A PRESENTATION AT THE 1980 ANNUAL MEETING OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION (April 7-11, 1980: Boston, Mass.)

ED191901

TM 800525

TABLE OF CONTENTS

Section I -- Correlations Between Students' Prior Subject Interest in the Course and Their Evaluations of Teaching Effectiveness in Two Different Universities
Section II -- A Path Analytic/Causal Model Relating Students' Prior Subject Interest and Other Student, Course and Instructors Characteristics to Students' Evaluations of Teaching Effectiveness
Section III -- Relationships Among Students' Evaluations of Teaching Effectiveness, Instructor Self Evaluations of Their Own Teaching Effectiveness, and Students' Prior Subject Interest

LIST OF TABLES

Table One -- Correlations Between Students' Prior Subject Interest and Their Evaluations of Teaching Effectiveness at Two Different Universities
Table Two -- Correlations Between 16 Background Variables and 11 Student Evaluation Scores, and the Unique Contribution of Each Background Variable
Table Three -- Path Analytic/Causal Model Relating Prior Subject Interest, Reason For Taking Course, Expected Grade and Woakload/Difficulty to Student Ratings
Table Four -- Correlations Among Student & Instructor Evaluations of Instructional Effectiveness and Student & Instructor Ratings of Students' Prior Subject Interest
Table Five -- Canonical Correlations Between Evaluations of Instructional Effectiveness and Ratings of Students' Prior Subject Interest

LIST OF APPENDICES

Appendix I -- Summary of Factor Analysis and Factor Reliability Coefficients of the Student Evaluation Instrument Used at UCLA
Appendix II -- Summary of Factor Analysis and Factor Reliability Coefficients of the Student Evaluation Instrument Used at USC
Appendix III -- Factor Analyses of Students' Evaluations of Teaching Effectiveness and the Corresponding Instructor Self Evaluations of Their Own Teaching
Appendix IV -- Correlations Between Student and Instructor Self Evaluations of Teaching Effectiveness: A Multitrait-Multimethod Matrix
Appendix V -- Complete Correlation Matrix of Student & Instructor Evaluations of Teaching Effectiveness: Variables Include the Nine Evaluation Factor Scores, Overall Course Rating, Overall Instructor Rating, and Students' Prior Subject Interest as Rated By Instructors and Students

Prior Subject Interest, Students' Evaluations, and Instructional Effectiveness

Student evaluations are designed to measure the effectiveness of teaching. Using a construct validation approach, researchers have sought to demonstrate that student ratings are positively related to variables identified as indicative of effective teaching and unrelated to variables that are not. Student ratings have been successfully validated by a variety of criteria, the most common one being performance on a standardized examination. When investigators examined different sections of the same course taught by different instructors, the results demonstrated that the sections that did best on the standardized examination given to all sections were also the ones who evaluated their instructors more favorably (Centra, 1977; Cohen & Berger, 1970; Frey, 1973, 1978; Marsh, Fleiner & Thomas, 1975; Marsh & Overall, in press). Student ratings have also been validated with other criteria including the ratings of former students (Centra, 1973; Marsh, 1977; Marsh & Overall, 1979a), affective course consequences such as feelings of subject mastery and a disposition to apply the subject matter further (Marsh & Overall, in press), and the self ratings of the faculty actually being evaluated (Doyle & Crichton, 1978; Webb & Nolan, 1955; Marsh & Overall, 1979b; Marsh, Overall & Kesler, 1979a). However, the validity coefficients reported in these studies did not approach unity and are generally in the range of .4 to .8. This means that while student ratings are related to the above validity criteria, much of their variance remains to be explained by other variables. The demonstration that student ratings are positively related to a variety of generally acceptable criteria of teaching effectiveness does not rule out the possibility that they are also related to extraneous factors.

Faculty generally do believe that student ratings are biased by a number of factors unrelated to the actual quality of teaching. Faculty at one of the institutions at which this study was conducted indicated that about half of a list of 17 "potential biases" actually would cause a substantial bias in student ratings (Marsh and Overall, 1979). Over half the respondents cited Course Difficulty (72%), Grading Leniency(68%), Instructor Popularity(63%), Student Interest in Subject Before Course(62%), Course Workload(60%), Class Size(55%), Reason For Taking the Course(55%), and Student's GPA(53%). Whether or not any one or a combination of these potential biases actually impact ratings of effective teaching in a way that is unrelated to the actual quality of teaching, faculty perceive that they do. As long as this perception persists, the utilization and interpretation of student evaluations will be limited.

Reviewing a broad spectrum of student evaluation literature, researchers (Costin, Greenough and Menges, 1971; Hildebrand, Wilson and Dienst, 1973; Marsh, 1978, in press; Menges, 1973; McKeachie, 1973; Remmers, 1963) have generally concluded that a variety of potential biases to student ratings of instruction are apparently of little consequence. However, many researchers have also reported that students' prior interest in the subject was related to student ratings reported at the end of the term. Feldman (1977), in an extensive review of the literature, found positive relationships, of substantial size in some studies, between subject interest and student ratings. Marsh(1978) reported that of the 15 background variables considered in his study,

Prior Subject Interest was most highly correlated with student ratings. Similarly, Burton (1975) found that enthusiasm for the subject was consistently more highly correlated with student ratings than any of the seven other background variables that she investigated. Consequently, while most student/course/instructor characteristics, either separately or in combination, do not seem to have a large impact on student ratings, Prior Subject Interest has been shown to have higher correlations with the ratings than most.

The purpose of this paper is to examine the relationship between students' Prior Subject Interest and: 1) students' evaluations of teaching effectiveness; 2) other student/course/instructor characteristics that might influence student ratings; 3) faculty self evaluations of their own teaching effectiveness. Student evaluations were collected in at two different universities, and faculty self evaluations of their own teaching were also collected at one of these settings. The results will be used to determine the answers to several different questions that will be examined in a separate section of this paper:

1. To what extent are student evaluations of teaching effectivenss correlated with students' Prior Subject Interest in the course?

2. What is the direct and indirect contribution of Prior Subject Interest when its effect is combined with that of other background/demographic variables commonly posited to influence student ratings?

3. What are the relationships between Prior Subject Interest, students' evaluations of teaching effectiveness, and instructor self evaluations of their own teaching effectiveness?

For example, one possible outcome would be to demonstrate that Prior Subject Interest is similarly related both to student and to instructor ratings of teaching effectiveness. This would suggest that Prior Subject Interest has an impact on the educational process that is reflected in the student ratings rather than being a bias specific to student evaluations.

SECTION I

Correlations Between Students' Prior Subject Interest in the Course and Their Evaluations of Teaching Effectiveness in Two Different University Settings

The purpose of this section is threefold: to determine 1) to what extent students' Prior Subject Interest is substantially correlated with student ratings of teaching effectiveness; 2) whether the correlation varies with different components of the student ratings; and 3) whether the findings generalize across different settings.

METHOD

The UCLA Setting. During the fall semester of the 1973 academic year, student evaluation forms were collected in 591 undergraduate classes at the University of California, Los Angeles. The student ratings were summarized by 10 evaluation scores; eight evaluation factor scores and two overall summary items. The eight evaluation factors identified by factor analysis were: Enthusiasm, Breadth, Organization, Interaction, Learning, Examinations, Assignments, and Workload/Difficulty (See Appendix I for a summary of the factor analysis.) Students were also asked to indicate what their subject interest in the course had been before the start of the semester. Howard and Schmeck (1979) have shown that estimates of pre-course interest collected at the end of the course show little difference from those actually collected prior to the course. Class average responses to the prior subject interest item were correlated with each of the student evaluation scores.

The USC Setting. During the 1976-77 academic year at the University of Southern California student evaluation of instruction surveys were administered in a total of 511 undergraduate courses taught by faculty in the Division of Social Sciences. Graduate level courses and courses taught by teaching assistants were not included in the analyses. Student evaluation instruments were sent to faculty in charge of all courses several weeks before the end of the semester and were actually used in virtually all of these courses. An average of 78% of the students enrolled in classes completed the survey forms.

The student ratings were summarized by eleven evaluation scores, nine evaluation factor scores and two overall summary items. These were similar to those described in the UCLA setting except that the Interaction factor was divided into two separate components--Group Interaction and Individual Rapport. (See Appendix II for a description of the factor analysis.) As in Study I students indicated their Prior Subject Interest as part of the evaluation instrument administered at the end of the term, and their class average responses were correlated with each of the student evaluation scores. All of the correlations were based upon class average responses.

RESULTS

The results of the UCLA study (See Table One) show that Prior Subject Interest was significantly correlated with four of the ten evaluation scores, including the two overall summary ratings. However, correlations with Prior Subject Interest were higher than .20 for only two of the evaluation scores; the Learning/Value (r=.39) factor score and the Overall Course Rating (r=.24). In the USC setting Prior Subject Interest was significantly correlated with six of the eleven evaluation scores. Two of these correlations were over .30 (Learning/Value--.44 and Overall Course Rating--.33). Four other correlations were between .20 and .30.

Insert Table One About Here

In both of these settings the pattern of results was similar, although the magnitude of the correlations were slightly higher in the USC study. Prior Subject Interest was correlated with several of the

evaluation scores in both settings, but the degree of correlation varied substantially for different dimensions. For both studies the correlations were highest for the Learning/Value dimension and for the Overall Course summary rating.

SECTION II

A Path Analytic/Causal Model Relating Students' Prior Subject Interest and Other Student, Course and Instructor Characteristics To Students' Evaluations of Teaching Effectiveness

The findings presented in Section I demonstrate that students' Prior Subject Interest is correlated with their evaluations of teaching effectiveness. However the interpretation of these findings is limited due to the correlational nature of the data.

The purpose of this analysis was to explore the causal nature of the relationship between student ratings of teaching effectiveness and a broad set of student, course, and instructor characteristics , including students' Prior Subject Interest. The results to be described are based upon the set of 511 USC courses that were discussed in Section I. The findings to be summarized are presented in more detail elsewhere (see Marsh & Overall, 1978; Marsh, in press). In this section a variety of multivariate techniques were applied to determine the combined effect of an entire set of background variables, the background variables that were most important, and how the combined effect varied for the different dimensions of the student evaluations. On the basis of these findings, four of the background variables -- including students' Prior Subject Interest -- were selected as being most important. These most important background variables were then used to develop and test a path analytic/causal model. The implications lead to a discussion of the importance of students' Prior Subject Interest in determining student ratings and its relationship in the context of the other background variables.

METHOD

Background Characteristics. The set of background characteristics consisted of 16 different variables describing the course, the students in the course, and the instructor. Selection of the set of variables was prompted by a review of the literature and the availability of information. The Workload/Difficulty variable has been included as both an evaluation score and a background variable. However, when the combined effect of the four most important variables was examined with path analysis, Workload/Difficulty was considered a background variable.

Two subsets of the background variables were given special attention in some of the analyses. The first subset was the Reason for Taking the Course. Students selected one of five possible reasons for taking the course, or left the item blank if not applicable. The percentages of students indicating each of the reasons were included as five separate variables. However, these five variables should be interpreted cautiously, since a higher percentage for any one suggests a

lower percentage for another. The second subset consisted of four highly correlated variables that described Class Level. The 16 background variables and brief descriptions are found in Appendix VI.

Procedures and Analysis.

Data analyzed in this section are as described in the USC setting in Section I. All analyses in this study were performed on class average responses for the 511 courses. Each of the 16 background variables was correlated with the student evaluation factors and two summary ratings.

Multiple regression was used to determine both the proportion of variance in the student ratings that could be uniquely explained by each of the background variables and to determine the smallest set of background variables that would best predict each of the evaluation scores. These analyses were used as a criterion for determining the most important background variables.

On the basis of this initial analysis, four background variables, including students' Prior Subject Interest , were selected as the most important in predicting the student evaluation scores. A path analytic/causal model was then developed to describe how these four combined in the prediction of the student ratings. The actual variables and the linkages between these variables, were based upon findings described in the results section. The path coefficients were standardized beta weights derived from a series of multiple regression analyses (see Kerlinger & Pedhazur, 1973; Blalock, 1971).

RESULTS

Bivariate Relationships

Linear correlations between each of the 16 background variables and the 11 evaluation scores are presented in Table Two. Correlations as small as r= .09 were statistically significant,although of little practical significance. Consequently, attention was focused on those relationships which accounted for at least 5% of the variance in any one of the evaluation scores (i.e., correlations of at least r = .23). Of the 175 correlations in Table Two only 18 met this criterion, and only 3 of these 18 accounted for as much as 10% of the variance; Prior Subject Interest was positively correlated with both Overall Course Rating and the Learning/ Value evaluation scores, and Enrollment was negatively correlated with quality of Group Interaction. None of the 16 background variables accounted for more than 5% of the variance in even a majority of the evaluation scores, and only three (Prior Subject Interest, Workload/Difficulty and Expected Grade) accounted for more than one evaluation score.

Insert Table Two About Here

Multivariate Relationships

Multiple regression was used to determine the unique contribution

of each of the individual background variables to each of the the evaluation scores (see table Two). The results of these analyses indicate that much of the variance accounted for by any one background variable was redundant with variance accounted for by others. Although 18 of 175 relationships between background variables and evaluation scores accounted for 5% of the variance, only seven relationships uniquely accounted for 5% (i.e., variance which was not also explained by other background variables). Only four background variables uniquely accounted for as much as 5% of the variance in even one of the evaluation scores: Prior Subject Interest, Workload/Difficulty, Expected Grade, and Reason For Taking the Course.

An unusual relationship consistently appeared between Expected Grade, Workload/Difficulty, and the evaluation scores that is contrary to the present body of literature in this area. Expected Grade and Workload/Difficulty both tended to be positively related to each of the evaluation scores, but were negatively related to each other ($r = -.29$). This event illustrates cooperative or reciprocal suppression, which is described by Cohen and Cohen (1975). While this occurrence is interesting, it also complicates interpretations. The combined effect of the two variables is necessarily greater than the sum of their individual effects. Furthermore, the supposedly "unique" variation attributed to either of these variables tends to be more than the variance explained before the effect of other variables has been removed.

Multiple regression was also used to determine the subset of background variables that best predicted each of the student evaluation scores. The set of 16 background variables was entered into a step-wise multiple regression to predict each of the 11 student evaluation scores. An additional variable was included only if it added at least 1% to the variance which had already been explained by the previous set of variables. The percentage of variance which could be explained in each of the different evaluation scores (corrected for the number of variables included in the regression equation) varied dramatically, ranging from from 0% for Organization to 25% for Learning/Value. Four of the background variables consistently appeared in the final regression equations: student ratings tended to be higher when 1) Prior Subject Interest was higher; 2) Workload/Difficulty was higher; 3) Expected Grades were higher; and 4) the percentage of students enrolling for "General Interest" (as opposed to Major Requirement, Minor/Related Field, or General Education Requirement) was higher.

Causal Modeling.

A difficult problem in multivariate research is the search for "the" most important independent variables (in this study, the background variables) to be used in predicting one or more dependent variables (the evaluation scores). Many alternative criteria with their strengths and weaknesses have been discussed elsewhere, but the most commonly suggested are: 1) simple correlations, 2) beta weights in the final regression equations, 3) the change in variance explained as each new variable enters a regression equation, and 4) the variance uniquely defined by each independent variable. Each of these criteria suggested the same three, or perhaps four, background variables were most important: Prior Subject Interest, Workload/Difficulty, Expected Grade,

and perhaps Reason For Taking the Course. These were the variables considered in the causal models.

The next step was to establish a temporal or causal ordering of these four most important background variables. On a temporal basis Prior Subject Interest and Reason for Taking the Course (represented by percent indicating "General Interest Only") were assumed to come first. These two variables were uncorrelated ($r = +.05$, not statistically significant). Consequently, there was no hypothesis that any causal linkage or ordering of these two variables and their combined effect was approximately the sum of the two variables' separate effects.

A temporal basis could not be used to determine the causal ordering of the other two background variables, Expected Grades and Workload/Difficulty. Three hypotheses are viable: 1) Expectations of lower grades might cause perceptions of greater Workload/Difficulty; 2) Higher levels of Workload/Difficulty might cause the expectations of higher grades ; and 3) these two influences can work simultaneously, canceling each other out. In fact, there was a substantial negative relationship between the two variables ($r = -.29$, $p<.001$), offering clear support for the first hypothesis. Consequently, the various causal models were based upon the assumption that Expected Grades preceded Workload/Difficulty in the causal ordering.

The proposed path analysis/causal model was based upon all four background variables (see Table Three). Due to the lack of correlation between Prior Subject Interest and Reason For Taking the Course, a linkage between these variables was not included in the model. Preliminary investigation had also demonstrated that the direct causal coefficient connecting Reason For Taking the Course and Expected Grade was almost zero ($P = .004$), and so this linkage was also excluded in the final model. Other linkages in the model were derived as described above. A preliminary test of the adequacy of the model was that the sum of the coefficients representing the direct causal linkage, the indirect causal linkage, and the spurious or noncausal linkage for each variable was very close to the correlation between that variable and the corresponding student rating. In neither case did the correlation derived from the model differ from the original correlation by more than .02, indicating that the few linkages that were excluded from the model had little impact.

Insert Table Three About Here

Inspection of Table Three indicates that an average of 12.8% (the sum of the four Total Causal Variance estimates) could be explained by the four background variables. This value is generally similar to that obtained with other multivariate techniques described earlier.

Averaged across all 10 evaluation scores, the Total Causal Variance estimate was highest for Prior Subject Interest (5.1%), followed by Workload/Difficulty (3.6%), Expected Course Grade (2.6%), and Reason For Taking Course (1.5%). Some of the Total Causal Variance attributed to Prior Subject Interest was indirect in that it worked through either Expected Grade or Workload/Difficulty. The reciprocal or

mutual suppressing effect of Expected Grade and Workload/Difficulty made their interpretation difficult; both these variables were positively related to Prior Subject Interest and the student ratings, but were negatively related to each other (path coefficient -.34). The direct causal effect between Expected Grades and student ratings was positive, but the indirect effect was negative. Consequently, the Total Causal Variance was actually less than the Direct Causal Variance.

The results in this section provide evidence for the effect of various student/course/instructor characteristics on student ratings of instructional effectiveness. Findings support the conclusion that the impact of Prior Subject Interest on student ratings accounts for more variance than any of the other background variables considered. These findings also indicate that students' entry level of subject interest has some impact on evaluations of instructional effectiveness.

Section III

Relationships Among Students' Evaluations of Teaching Effectiveness, Instructor Self Evaluations of Their Own Teaching Effectiveness, and Students' Prior Subject Interest

The findings presented earlier demonstrate that students' Prior Subject Interest is related to students' evaluations of teaching effectiveness, and that it has a greater impact than other background influences typically considered. The results of the path analytic/causal model even suggest that the student ratings may be "caused" by Prior Subject Interest, though this interpretation was limited by the correlational nature of the data. However, in each of the studies student ratings of teaching effectiveness was the only criterion considered. The question as to the possible relationship between students' Prior Subject Interest and underlying dimensions of effective teaching that are measured by other methods was unanswered. While researchers have clearly demonstrated that student ratings are related to a variety of different criteria of effective teaching, the ratings generally account for 25% or less of the criterion variance. Consequently, it is possible that students' Prior Subject Interest is related to students' ratings of teaching effectiveness, but not to any other criteria of effective teaching. This is considered a possibility since both the measures of students' Prior Subject Interest and teaching effectiveness are based upon ratings from the same individuals. A more persuasive demonstration would consist of showing that students' Prior Subject Interest -- measured by at least two different methods -- is correlated with more than one distinct measure of teaching effectiveness.

Marsh (Marsh, Overall & Kesler, 1979b; Marsh & Overall, 1979a) asked instructors to evaluate their own teaching with the same evaluation form that was used by their students. Both student and instructor self evaluations were reliable, and the same set of nine evaluation factors previously discussed were shown to underlie both sets of ratings. Student and instructor ratings showed agreement in that their ratings were significantly correlated (median r = .45 in one study and .49 in the other) and there was little difference in the mean responses given by the two groups. Furthermore, student-instructor

agreement on each of the different factors was independent of agreement on other factors; while correlations between student and instructor ratings on the same factors were high (median r = .45), correlations between their ratings on different factors was low (median r = 0.0). For example, instructor self evaluations of Enthusiasm correlate highly with students' evaluations of Enthusiasm (r = .54), but not with any of the other student rating factors (highest r less than .15). These results argue that student ratings do measure a general construct of effective teaching and also argue for the distinctiveness of the different evaluation factors.

As part of this same study, both students and instructors were asked to evaluate students' Prior Subject Interest. The purpose of the present analysis is to determine if Prior Subject Interest, as measured by both student and instructor ratings, is similarly correlated with student and instructor ratings of teaching effectiveness. If this multivariate relationship is demonstrated, then it can be argued that students' Prior Subject Interest is a variable that actually impacts on the quality of teaching, not just on student ratings. However, if the effect of Prior Subject Interest is specific to the student ratings and does not generalize to instructor self evaluations, alternative explanations would be necessary.

METHOD

During the academic year 1977-78 student evaluations were collected in virtually all courses offered in the Division of Social Sciences at the University of Southern California. Evaluations were administered shortly before the end of the term, generally by a designated student in the class or by a staff person. Students were told that the evaluations would provide feedback to instructors, and would be considered as part of personnel decisons. The surveys were completed by an average of 76% of the students enrolled in each class.

The evaluation instrument (see Appendix II) consisted of 35 evaluation items adapted from Hildebrand, Wilson & Dienst(1971) and Marsh, Overall & Thomas (1976). The median reliability of individual evaluation items--intraclass correlation coefficients based upon sets of responses from 25 students per class was .88(see Appendix II). A factor analysis (see Appendix III) of the student ratings of all undergraduate courses taught by regular faculty revealed nine separate evaluation factors. The reliability of the factors, coefficient alphas, varied from .88 to .97 (see Appendix II).

Instructor self evaluation surveys were sent to all teachers who had been evaluated by students in at least two different courses during the same term. Instructors were asked to evaluate the effectiveness of their own teaching in both courses. These surveys were completed after the end of the term, but before summaries of the student evaluations were returned. Instructors evaluated both courses with a set of items identical to those used by students, except for slight modification in the wording. They were specifically instructed to rate their own teaching effectiveness and not to report how students would rate them. A total of 181 (78%) surveys were returned. Since faculty had been requested to rate the effectiveness of their teaching in both classes they taught, self ratings for a total of 331 courses were completed, 183 undergraduate courses taught by faculty, 45 graduate level courses, and

.03 courses taught by teaching assistants.

Eleven evaluation scores--factor scores representing the nine evaluation factors and overall ratings of the teacher and the course were used to summarize the student ratings and the instructor self ratings. Evaluation factor scores were weighted averages of standardized responses to each evaluation item. The weights, factor score coefficients, were derived from the factor analysis described in Appendix II.

RESULTS

Relationship Between PSI and Separate Evaluation Scores.

The set of student evaluations were correlated with four different combinations of the two Prior Subject Interest (PSI) ratings:

1) Student ratings of their own Prior Subject Interest
2) Instructor ratings of students' Prior Subject Interest
3) An unweighted average of the two Prior Subject Interest ratings (after each was standardized)
4) A "best weighted" average of the two Prior Subject Interest ratings -- the multiple R resulting from a multiple regression in which both Prior Subject Interest ratings were used to predict each of the student ratings.

The results of these correlations are presented in Table Four. Although the correlations between instructor ratings of students' Prior Subject Interest and student ratings are lower, there is a consistent pattern in the four sets of correlations. For each set of correlations, the highest correlation is with the Learning/Value factor, the second highest is with the Overall Course Rating, and statistically insignificant correlations occur with the Organization, Individual Rapport, and Examination factors.

Insert Table Four About Here

Correlations with the unweighted average of the two Prior Subject Interest ratings are consistently higher than correlations with either of the ratings separately. Furthermore, the "best weighted" average of the two Prior Subject Interest variables produces correlations that are nearly the same as those resulting from the unweighted average. This occurs because the best weights are approximately equal weighting.

Instructor self evaluations of their own teaching effectiveness were also correlated with each of the four combinations of the Prior Subject Interest ratings (see Table Four). Although the magnitude of these correlations is somewhat lower than was observed with the student ratings (perhaps due to the lower reliability of the instructor ratings), the pattern of findings is quite similar to that just described for the student ratings. The Learning/Value and Overall Course ratings were best predicted, correlations with the Organization, Individual Rapport, and Examination factors were statistically

insignificant, and the "best weighting" of the two Prior Subject Interest ratings was again an approximately equal weighting. It is interesting to note that while instructor self evaluations were generally somewhat less correlated with the Prior Subject Interest ratings, their ratings of their Enthusiasm was consistently more highly correlated; i.e., instructors see their own enthusiasm as more closely related to students' Prior Subject Interest than do students.

In the final set of correlations, combined measures of the two Prior Subject Interest ratings were correlated with combined measures of both student and instructor evaluations. In the first of two analyses, an unweighted average of the two Prior Subject Interest ratings was correlated with an unweighted average of student and instructor evaluations. One interpretation of this analysis is that each of the instuctional evaluation scores represents a hypothetical or underlying dimension that is equally well measured by both student and instructor ratings of that dimension. Similarly, Prior Subject Interest is an underlying dimension that is equally well measured by the student and instructor ratings of that variable. The pattern of correlations is again quite similar to that described for both of the student and the instructor ratings considered separately; correlations are highest for the Learning/Value and Overall Course evaluations, and correlations with the Organization, Individual Rapport, and Examination factors do not reach statistical significance. A particularily interesting fact is that the correlations with both the Learning/Value and Overall Course ratings are higher than any of those previously observed in the table. This provides strong support for the existence of the underlying constructs that were hypothesized.

In the analysis just described, student and instructor evaluations of teaching effectiveness were equally weighted in determining their average, as were student and instructor ratings of students' Prior Subject Interest. Canonical correlation provides an alternative analysis in which the "best weighted" (as opposed to equally weighted) average of the instructional evaluations is related to the "best weighted" average of the Prior Subject Interest ratings.

Results of a canonical correlation analyses need to be interpreted cautiously. For example, in the present study, if a random variable were added to each case and this random variable were included with both the Prior Subject Interest variables and the instructional evaluation variables, the resulting canonical correlation would be 1.0. The random variable would be perfectly correlated with itself and the best weightings would consist of a 1.0 weighting of the random variable in each set and a zero weighting for all the other variables. Consequently, while the resulting canonical correlation would be 1.0, it would say nothing of how quality of teaching is related to Prior Subject Interest.

Cooley and Lohnes (1971; 1976) describe a redundancy measure which indicates how much variance in one set of variables can be predicted by a second set of variables. Conceptually, the interpretation of this redundancy statistic is similar to r squared or multiple regression. It should be noted, however, that canonical correlation maximizes the canonical correlation and not the redundancy statistic, and these two are not necessarily even monotonically related.

In the last column of Table Four, the canonical correlation (and the redundancy statistic) relating the pair of Prior Subject Interest ratings to each pair of instructional evaluations (i.e., student and instructor ratings on the same dimension) is presented. The same pattern of results described for each of the other sets of correlations is again evident. The strongest relationships (both in terms of canonical R and redundancy) occur with the Learning/Value and Overall Course ratings, while those for the Organization, Rapport, and Examination factors are lowest. Paradoxically, the "variance explained" (the redundancy statistic) in this analysis is generally less than that observed with the correlations resulting from the unweighted averages. This is a consequence of the problem with interpretation of canonical correlation noted above.

Relationship Between PSI and the Combined Set of Evaluations.

The analyses presented above determined the relationship between students' Prior Subject Interest and each of the instructional evaluation variables considered separately. Analyses presented in this section determine the relationship between Prior Subject Interest and the combined set of instructional evaluations. The analyses are performed on only the nine evaluation factors, since the two overall rating items were included in the factor analysis used to determine the factors (see Appendix II & III).

The results of three different canonical correlations (canonical R's, redundancy, and the loadings on each variable) are presented in Table Five. In these three analyses, the pair of Prior Subject Interest variables were correlated with 1) the set of 9 student evaluations factors, 2) the set of 9 instructor self evaluations, and 3) the combined set of 18 student and instructor evaluations.

Insert Table Five About Here

The two Prior Subject Interest variables accounted for 3.0% of the variance in the instructor ratings, 4.7% of the variance in the student ratings, and 3.7% of the variance in the combined set of student and instructor ratings. In predicting the instructor self ratings, the instructors' rating of students' Prior Subject Interest was weighted somewhat more heavily; students' ratings of their own Prior Subject Interest was weighted somewhat more heavily in predicting student ratings. However, the two Prior Subject Interest variables were almost equally weighted in predicting the combined set of student and instructor evaluations. In each of these analyses, the Learning/Value factor was best predicted, while the loadings for the Organization, Individual Rapport, and Examination factors were most poorly predicted. In each of the analyses the Assignments and Workload/Difficulty factors also had substantial loadings. Instructor self evaluations of their own Enthusiasm were more strongly related to the Prior Subject Interest variables than were student evalutions of Enthusiasm, while student ratings of Group Interaction were weighted more heavily than were instructor evaluations of this variable.

The results discussed in this section present a very consistent

picture. Students' Prior Subject Interest, as perceived by both the instructors and the students themselves, is correlated with both student and instructor evaluations of teaching effectiveness. Across the entire set of evaluation dimensions the magnitude of this relationship is reasonably small -- less than 5% of the variance in the evaluations is predictable in terms of Prior Subject Interest. However, the strength of the relationship varies dramatically for the different dimensions of the evaluations. In each of a variety of bivariate and multivariate analyses the Prior Subject Interest ratings were most highly correlated with the Learning/Value factor and the Overall Course rating (an individual item that loads highly on the Learning/Value factor). The pattern of the findings for both student and instructor evaluations was found in each of the analyses.

DISCUSSION & IMPLICATIONS

In each of the three studies, evaluations of instructional effectiveness were significantly correlated with students' Prior Subject Interest. Furthermore, there was also substantial correspondence in the specific evaluation factors that were and were not related to Prior Subject Interest. Ratings of Learning/Value and the Overall Course Rating were most highly related to Prior Subject Interest, while Breadth of Coverage, and Organization were correlated lower with it (or uncorrelated). Prior Subject Interest was shown to be more highly related to student ratings, though it generally accounted for an average of only about 5% of the variance, than any other background variable considered. Furthermore controlling for the effect of Prior Subject Interest reduced the relationship between other background variables (particularly Expected Grades) and student ratings. However, the three studies taken together argue that Prior Subject Interest can appropriately be considered a variable that impacts on the quality of the educational process and not an influence or bias that is specific to student ratings. Both student and faculty perceptions of students' Prior Subject Interest were positively correlated with both student and faculty self ratings of instructional effectiveness. Higher student interest in the subject apparently creates a more favorable learning environment and facilitates effective teaching. This argues that Prior Subject Interest actually affects the quality of teaching and that this effect is reflected in this study of student ratings as well as faculty self evaluations.

TABLE ONE

Correlations Between Students' Prior Subject Interest in Course and Their Evaluations of Teaching Effectiveness At Two Different Universities

	UCLA 1973 N=591 classes	USC 1976 N=511 classes
Learning/Value	0.39**	0.44**
Enthusiasm	0.03	0.23**
Organization	0.06	-0.03
Group Interaction	----	0.29**
Individual Rapport	0.09	0.09
Breadth of Coverage	-0.05	-0.03
Examinations	0.17**	0.03
Assignment	0.09	0.23**
Workload/Difficulty	0.12	-0.12
Overall Course Rating	0.24**	0.33**
Overall Instructor Rating	0.13*	0.20**

* $p<.05$; ** $p<.01$

NOTE: Evaluation instruments used at the two universities were similar, but not the same. The UCLA form had only one Interaction factor that combined aspects of both the Group Interaction and Individual Rapport factors on the USC form. Factor analyses of both forms are included in Appendices I and II.

TABLE TWO

Correlations Between 16 Background Variables and 11 Student Evaluation Scores, and the Unique Contribution of Each Background Variable
(N=511 Averages)

BACKGROUND VARIABLES	STUDENT EVALUATION SCORES										
	Overall Course	Overall Instructor	Learning	Enthusiasm	Organ-ization	Group Interaction	Individual Rapport	Breadth	Exam	Assignments	Workload/ Difficulty
Prior Subject Interest	33 (16)	20 (08)	44 (26)	23 (14)	-03 (-03)	29 (09)	09 (00)	-03 (-09)	03 (-05)	23 (18)	12 (17)
Workload/Difficulty	23 (30)	14 (19)	12 (19)	06 (11)	01 (04)	-02 (00)	01 (03)	15 (14)	10 (16)	23[a] (24)	xx xxx
Expected Grade	21 (24)	20 (21)	29 (23)	20 (18)	01 (03)	31 (19)	17 (14)	-02 (01)	18 (21)	13 (14)	-29 (-36)
Reason for Taking Course	(18)	(13)	(18)	(04)	(18)	(10)	(08)	(24)	(10)	(20)	(14)
% General Interest Only	16 (18)	12 (14)	15 (15)	09 (09)	16 (19)	07 (10)	-02 (00)	19[a] (19)	10 (14)	18 (20)	-13 (07)
% Major Elective	16 (03)	13 (00)	26 (09)	06 (-08)	-03 (00)	21 (00)	04 (-00)	18[a] (13)	02 (00)	15 (03)	-06 (10)
% Major Requirement	-15 (-15)	-12 (-13)	-18 (-16)	-07 (-08)	-08 (-14)	-04 (03)	01 (00)	26[a] (23)	-02 (-07)	-17 (-16	17 (15)
% General Ed Require	-11 (00)	-08 (03)	-17 (-13)	-04 (04)	03 (07)	-28 (-04)	-06 (09)	03 (00)	-09 (00)	-06 (09)	-12 (-08)
% Minor/Related Field	07 (06)	06 (04)	07 (05)	03 (00)	02 (00)	02 (00)	04 (00)	12 (10)	01 (00)	07 (05)	-01 (00)
Class Level	(00)	(00)	(00)	(00)	(13)	(00)	(12)	(11)	(11)	(04)	(10)
Course Level	17 (00)	14 (00)	21 (00)	12 (04)	-08 (-08)	29 (03)	14 (00)	13 (04)	04 (-03)	11 (00)	06 (03)
% Frosh-Soph in Class	-12 (00)	-12 (00)	-18 (00)	-08 (-03)	-01 (00)	-28 (00)	-17 (-11)	01 (00)	-09 (00)	-05 (00)	-03 (-03)
% Jr.-Sr. in Class	11 (00)	01 (00)	07 (00)	21 (00)	04 (03)	05 (00)	13 (00)	03 (00)	10 (07)	09 (-03)	-01 (-05)
Avg. "Year in School"	11 (00)	10 (00)	20 (00)	11 (05)	-06 (00)	27 (00)	19 (11)	-06 (-03)	03 (-07)	04 (00)	03 (00)
Overall GPA (prior)	07 (-10)	07 (-06)	10 (-10)	07 (-05)	-06 (00)	17 (00)	14 (00)	04 (00)	07 (00)	13 (00)	12 (19)
% Division Majors	15 (03)	14 (07)	15 (00)	03 (00)	05 (09)	29 (10)	08 (00)	08 (03)	13 (06)	11 (00)	15 (09)
Enrollment	-10 (00)	-09 (00)	-14[a] (00)	01 (05)	-03 (00)	-32[a] (-16)	-18 (-10)	01 (00)	-13 (-04)	-04 (-04)	01 (00)
Teacher Rank	-02 (-08)	-08 (00)	-10 (-00)	-12 (-10)	-10 (-09)	-14 (-10)	-05 (00)	24[a] (13)	-14 (-14)	13 (09)	11 (08)
MULTIPLE R^2 (% Variance Explained)	20%	8.9%	24.7%	5.9%	0%	23.0%	3.5%	11.3%	8.3%	12.3%	19.6%

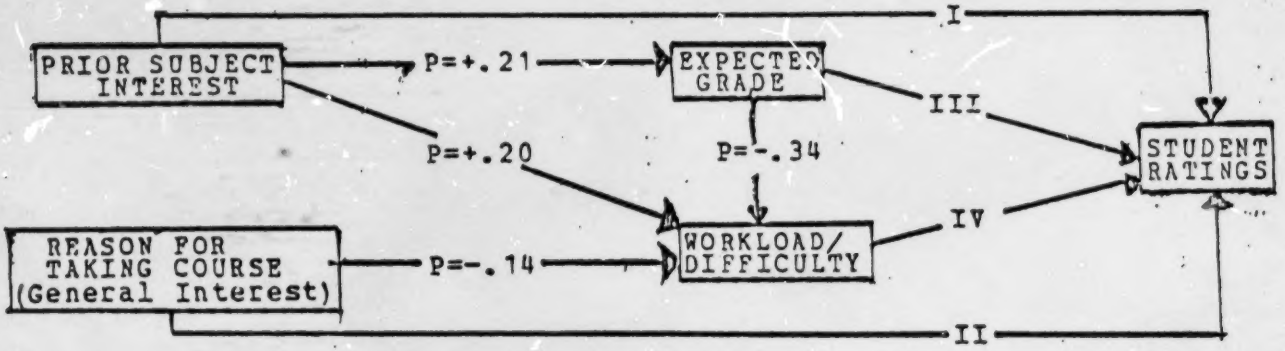
Note: Simple correlations, the values not in parentheses, are presented without decimal points. Correlations which are underlined indicate background variables accounting for at least 5% of the variance in an evaluation score.

Note: Uniqueness coefficients or the semipartial correlation coefficients, the values in parentheses, are the square root of the change in adjusted R^2 due to the introduction of each background variable after all other background variables have been entered. For the two subsets of background variables uniqueness was determined in two different ways. First, the square root of the change in R^2 was determined when all the variables in the subset were entered at once. Second, the square root of the change in R^2 was determined by entering each variable in the subset separately without considering the other variables in the subset. The uniqueness of any one variable in the subset may be as large or larger than that of the entire set since the R^2 was adjusted for the total number of variables in the regression equation.

[a]These relationships show substantial non-linear trends (i.e., quadratic and/or cubic components add at least 1% to the Variance Explained by the linear relationship and the Total Variance Explained by all components was at least 5%).

TABLE THREE

PATH ANALYSIS MODEL RELATING PRIOR SUBJECT INTEREST, REASON FOR TAKING COURSE, EXPECTED GRADE AND WORKLOAD/DIFFICULTY TO STUDENT RATINGS



	I PRIOR SUBJECT INTEREST			II REASON (General Interest Only)			III EXPECTED COURSE GRADE			IV WORKLOAD/ DIFFICULTY		
EVALUATION SCORES	Dir Caus	Tot Caus	Orig r	Dir Caus	Tot Caus	Orig r	Dir Caus	Tot Caus	Orig r	Dir Caus	Tot Caus	Orig r
Learning/Value	+36	+44	+44	+15	+13	+15	+26	+20	+29	+17	+17	+12
Enthusiasm	+17	+23	+23	+09	+08	+09	+20	+16	+20	+11	+11	+06
Organization	-04	-04	-03	+16	+16	+16	+03	+02	+01	+04	+04	00
Group Interaction	+21	+28	+29	+06	+06	+07	+30	+27	+31	+06	+06	-02
Individ Rapport	-05	+09	+09	-01	-02	-02	+18	+16	+17	+06	+06	+01
Breadth	-07	-03	-03	+23	+19	+19	+06	-01	-02	+21	+21	+15
Exams/Grading	-05	+03	+03	+12	+10	+10	+25	+18	+18	+20	+20	+10
Assignments	+11	+19	+20	+21	+17	+18	+19	+09	+13	+30	+30	+23
OVERALL COURSE	+23	+32	+33	+19	+15	+16	+26	+15	+22	+30	+30	+23
OVERALL INSTRUCTOR	+12	+20	+20	+13	+11	+12	+24	+17	+20	+17	+17	+10
Variance Components[a]	2.9%	5.1%	5.3%	2.3%	1.5%	1.8%	4.5%	2.6%	4.0%	3.6%	3.6%	1.8

NOTE: The method of calculating the path coefficients (p values in diagram), Direct Causal Coeficients (Dir Caus), and Total Causal Coefficients (Tot Caus) were described earlier.

a--The "Variance Component" was calculated by summing the squared coefficients, dividing by the number of coefficients, and multiplying by 100%.

TABLE FOUR

Correlations Among Student and Instructor Ratings of Instructional Effectiveness and Student and Instructor Ratings of Students' Prior Subject Interest (PSI)

	STUDENT EVALS CORRELATED WITH:				INSTRUCTOR SELF EVALS CORRELATED WITH:				AVERAGE OF STUDENT & INSTRUCTOR EVALS CORRELATED WITH:	
EVALUATION FACTORS	Student Ratings of their own PSI	Instructor Ratings of Students' PSI	Unwghted Average of both PSIs	Weighted Average of both PSIs	Student Ratings of their own PSI	Instructor Ratings of Students' PSI	Unwghted Average of both PSIs	Weighted Average of both PSIs	Unweighted Average of both PSIs	Weighted Average of both PSIs
	r =	r =	r =	Mult R =	r =	r =	r =	Mult R =	r =	Canon R(redun)
Learning/Value	.38**	.29**	.41**	.41**	.26**	.37**	.38**	.39**	.46**	.47(15.4%)**
Enthusiasm	.13*	.10	.14*	.14*	.15**	.18**	.20**	.20**	.20**	.20(2.9%)**
Organization	.06	.05	.07	.07	-.08	-.09	.10	.11	-.02	.14(0.8%)
Group Interaction	.27**	.14*	.25**	.27**	.03	.14	.10	.14	.20**	.27(4.5%)**
Individual Rapport	.06	.08	.08	.09	-.03	-.09	.08	.10	.00	.14(0.7%)**
Breadth of Coverage	.21**	.16*	.22**	.23**	-.05	.06	.01	.10	.14*	.24(2.6%)**
Examinations	.05	.01	.04	.05	-.03	-.10	-.08	.10	-.03	.12(0.7%)
Assignment	.25**	.17**	.25**	.26**	.18*	.06	.15*	.17**	.24**	.26(4.2%)**
Workload/Difficulty	.19**	.07	.15*	.19**	.16*	.08	.14*	.16*	.18**	.20(1.4%)**
SUMMARY ITEMS										
Overall Course	.30**	.23**	.32**	.33**	.21**	.26**	.30**	.30**	.37**	.38(9.1%)**
Overall Instructor	.21**	.16*	.23**	.23**	.18*	.13*	.17*	.19**	.25**	.26(4.5%)**
MEAN VARIANCE EXPLAINED[1](across all 11 evaluation scores)	4.73%	2.37%	5.02%	5.40%	2.15%	2.85%	3.44%	3.95%	5.47%	4.25%

* p<.05; ** p<.01

NOTE: Students' ratings (N=329 class averages) were correlated with their own ratings of their Prior Subject Interest(PSI), their instructor's rating of their PSI, and the unweighted average of the two PSI ratings. Multiple regression was then used to determine the best weighted average of the two PSIs in predicting the students' ratings. Similar analyses were performed on the instructor self evaluations. An unweighted average of the two instructional effectiveness ratings was then correlated with an unweighted average of the two PSI ratings. Finally, canonical correlation was used two determine the highest possible correlation between a weighted average of the two instructional effectiveness ratings and the two PSI ratings; the redundancy statistic (redun) -- see Cooley & Lohnes, 1976 -- is the percentage of variance in the two instructional effectiveness ratings (student and instructor ratings of the same factor) that can be explained by the two PSI ratings.

1 -- MEAN VARIANCE EXPLAINED was defined as the average of the squared correlation coefficents for all but the one cononical correlation analysis; in this instance it was the average of the redundancy statistics.

TABLE FIVE

Canonical Correlations Between Ratings of Instructional Effectiveness and Ratings of Students' Prior Subject Interest (PSI)

	Instructor Ratings	Student Ratings	Stud & Instr Ratings
Instructor Ratings of Learning/Value	0.827	---	0.644
Instructor Ratings of Enthusiasm	0.415	---	0.334
Instructor Ratings of Organization	-0.207	---	-0.178
Instructor Ratings of Group Interaction	0.157	---	0.172
Instructor Ratings of Individual Rapport	-0.139	---	-0.140
Instructor Ratings of Breadth of Coverage	-0.031	---	0.017
Instructor Ratings of Examinations	-0.140	---	-0.143
Instructor Ratings of Assignment	0.346	---	0.247
Instructor Ratings of Workload/Difficulty	0.325	---	0.238
Student Ratings of Learning/Value	-----	0.802	0.699
Student Ratings of Enthusiasm	-----	0.276	0.241
Student Ratings of Organization	-----	0.133	0.116
Student Ratings of Group Interaction	-----	0.516	0.426
Student Ratings of Individual Rapport	-----	0.143	0.139
Student Ratings of Breadth of Coverage	-----	0.443	0.386
Student Ratings of Examinations	-----	0.093	0.073
Student Ratings of Assignment	-----	0.509	0.438
Student Ratings of Workload/Difficulty	-----	0.341	0.268
Average of Squared Coefficients	0.1321	0.1782	0.1069
Instructor ratings of students' PSI	0.930	0.661	0.816
Student ratings of their own PSI	0.686	0.942	0.839
Canonical Correlation	0.473**	0.516**	0.584**
Redundancy in Evaluations Explained by Prior Subject Interest (PSI)	2.96%	4.74%	3.65%

** $p < .001$

NOTE: Three canonical correlations were performed relating the set of two Prior Subject Interest ratings (PSI) with Instructor ratings, with Student ratings, and with both Student and Instructor ratings. The redundancy statistic is an estimate of the percentage of variance in the set of instructional evaluations explained by the set of PSI ratings (see Cooley & Lohnes, 1976).

APPENDIX I

Factor Analysis of Student Evaluation Instrument (N=591 Class Average Response)

Factor Pattern Loadings

Evaluation Items (paraphrased)	Enthusiasm Concern	Breadth	Organization	Interaction	Learning	Examinations	Assignments	Difficulty
Enthusiastic about teaching	**.55**	.10	.12	.12	.12	.05	.13	- .04
Subject made more understandable	**.49**	.11	.29	.03	.10	.02	.18	- .17
Concern that students learn	**.38**	.03	.21	.31	.13	.22	.03	- .12
Course made more relevent	**.42**	.30	.09	.10	.19	.03	.10	- .16
Effectiveness of lecture/presentation	**.55**	.10	.22	- .11	- .04	.07	.26	- .05
Presented different points of view	.00	**.68**	.07	.24	.06	.04	.04	- .08
Presented background concepts	.17	**.67**	.11	- .08	.08	.01	.08	- .09
Contrasted implications of theories	.08	**.80**	.06	.03	.01	.04	.12	- .10
Presentations were well prepared	.32	.11	**.57**	- .06	.03	.09	.14	- .03
Materials prepared and explained	.17	.10	**.65**	.01	.06	.12	.15	- .07
Objectives stated and persued	.03	.06	**.68**	.06	.13	.12	.16	- .01
Workload evenly distributed	- .05	.02	**.36**	.10	.02	.15	.23	- .32
Students felt welcome	.11	.01	.16	**.63**	.07	.17	.00	- .12
Encouraged to question	.20	.05	.06	**.67**	.09	.10	.09	- .13
Encouraged to express own ideas	.01	.24	- .02	**.70**	.09	.08	.09	- .09
Effectiveness of class discussion	.25	.00	- .10	**.40**	.04	.13	.26	- .05
Learned something valuable	.14	.04	.16	.12	**.59**	.08	.14	- .06
Interest increased by course	.17	.12	.23	.10	**.63**	.06	.18	- .04
Intellectually stimulating	.12	.01	- .04	.01	**.69**	.11	.16	- .07
Mastery	- .02	- .02	.05	- .06	**.35**	.27	.21	- .24
Tested actual content	.04	.12	.16	.13	.23	**.50**	.03	- .17
Effectiveness of graded materials/exams	.01	.02	.02	- .07	.01	**.85**	.16	- .07
Exams/graded material feedback valuable	.05	- .04	.05	.15	- .08	**.65**	.16	.02
Readings/text valuable	- .01	.03	.02	- .12	.06	- .02	**.68**	.00
Assignments contributed to understand.	- .07	.00	.03	.05	- .02	.19	**.75**	- .03
Difficulty	.12	.04	.05	.07	.08	.03	.08	**.93**
Workload	.03	.02	.12	- .03	.03	.05	- .06	**.63**
Overall instructor	.42	.13	.23	.18	.17	.12	.12	- .12
Overall course	.18	.10	.15	.04	.49	.06	.29	- .07

NOTE: Loadings for variables designed to measure each factor (in bold boxes) are substantially higher than other loadings, confirming the factor pattern hypothesized. Factor Analysis was oblique with Delta Factor = -2.0 (Nie, et. al, 1976)

Factor Analysis of Student Evaluation Instrument (N=511 Class Average Responses)

Evaluation Items (paraphrased)	Mean	Standard Deviation	Factor Pattern Loadings I	II	III	IV	V	VI	VII	VIII	IX
I. LEARNING/VALUE											
Increased Interest as Course Consequence	3.91	0.56	69	14	-04	06	04	09	06	17	00
Learned Something Valuable	4.15	0.48	59	06	11	12	00	04	11	15	18
Learned & Understood Subject Matter	4.01	0.41	53	11	17	09	06	-09	10	12	-28
OVERALL COURSE RATING	3.83	0.65	44	23	12	07	05	07	20	17	10
Intellectually Challenging/Stimulating	3.98	0.54	43	17	03	08	-01	18	19	13	31
II. ENTHUSIASM											
Dynamic & Energetic	3.90	0.65	08	67	15	07	04	09	09	11	07
Enhanced with humor	3.85	0.65	01	67	16	00	08	06	07	10	00
Held your interest	3.66	0.67	14	65	26	06	02	03	03	10	01
Enthusiatic about Teaching	4.18	0.57	10	48	19	07	13	14	13	09	06
OVERALL INSTRUCTOR RATING	3.97	0.65	14	43	25	10	17	11	12	08	05
III. ORGANIZATION											
Materials Prepared & Explained	3.90	0.56	12	-06	70	03	07	14	14	10	04
Instructor Explanations Clear	3.90	0.56	18	12	57	14	03	08	12	08	-07
Lectures Facilitated Note Taking	3.77	0.62	08	-02	51	-19	06	27	08	11	-03
Objectives stated & pursued	3.94	0.53	20	-10	49	06	08	12	24	11	04
IV. GROUP INTERACTION											
Students shared ideas/knowledge	4.07	.59	08	10	-01	81	07	04	07	08	00
Encouraged to Participate	4.05	.60	11	12	03	80	07	00	00	07	00
Encouraged to Express Own Ideas	4.09	.55	06	12	04	73	16	07	11	04	00
Encouraged to Question & Given Answers	4.08	.55	09	13	17	62	16	04	12	07	-02
V. INDIVIDUAL RAPPORT											
Welcomed Seeking Help/Advice	4.13	.54	08	10	05	06	82	-02	10	03	-01
Interested in Individual Students	4.02	.57	06	18	06	17	69	-06	14	03	00
Accessible to Students	3.91	.56	-02	08	03	01	65	24	11	11	07
Friendly Towards Students	4.28	.49	00	-25	12	18	61	-08	10	06	-09
VI. BREADTH OF COVERAGE											
Presented Background of Concepts	3.97	.48	12	05	12	02	05	68	07	12	-03
Contrasted Implications	3.94	.49	06	10	03	07	08	67	01	20	04
Presented Different Points of View	4.03	.44	03	08	17	12	03	60	12	12	-03
Discussed Current Developments	4.14	.49	19	12	15	16	02	32	17	12	-06
VII. EXAMINATIONS											
Eval Methods Fair/Appropriate	3.80	.56	03	04	02	05	16	05	72	15	-06
Tested Actual Content	3.88	.55	09	02	10	02	06	09	67	14	-04
Exam Feedback Valuable	3.67	.59	03	05	09	10	16	-02	66	07	09
VIII. ASSIGNMENTS											
Readings/Text Valuable	3.72	.59	-02	-05	02	00	04	11	-01	91	02
Contributed to Understanding	3.88	.54	09	01	06	10	02	01	16	70	03
IX. WORKLOAD/DIFFICULTY											
Workload (Light-Heavy)	3.37	.61	10	02	02	08	00	00	00	00	09
Difficulty (Easy-Hard)	3.45	.52	-02	02	00	-01	-01	11	07	08	85
Hours Out of Class	2.61	.61	13	03	-10	01	10	07	-09	12	76
Pace (Too Slow-Too Fast)	3.09	.39	-09	01	11	-10	-05	12	14	04	60

1--Factor Analysis was Oblique (correlated) with the Delta Factor=-2.0 (Nie, et. al, 1976)

2--First nine eigenvalues were 19.0, 3.3, 2.3, 1.5, 1.2, 1.0, .76, .60, .50

3--Correlations between Factors ranged from r=-.01 to r=.49 (Median r=.27)

4--All items except Workload/Difficulty were answered along 5-point response scale (1-Very Poor, 3-Moderate 5-Very Good). Workload/Difficulty items varied on 5-point response scale with end-points above, except for Hours (1-0 to 2, 2-2 to 5, 3-5 to 7, 4-8 to 12.5-Over 12).

APPENDIX III

Factor Analyses of Students' Evaluations of Teaching Effectiveness and the Corresponding Faculty Self Evaluations of Their Own Teaching in All 329 Courses

Evaluation Items (paraphrased)	Factor Pattern Loadings								
	I	II	III	IV	V	VI	VII	VIII	IX
I LEARNING/VALUE									
Course Challenging/Stimulating	42 (40)	23 (25)	09 (-10)	04 (04)	00 (-03)	15 (27)	09 (05)	16 (23)	29 (20)
Learned Something Valuable	53 (77)	15 (02)	10 (-02)	09 (04)	01 (01)	10 (00)	10 (04)	17 (09)	16 (06)
Increased Subject Interest	57 (70)	12 (05)	08 (07)	08 (07)	02 (-03)	18 (08)	03 (-04)	19 (05)	14 (-02)
Learned/Understood Subject Matter	55 (52)	12 (12)	13 (12)	05 (03)	03 (11)	02 (-01)	19 (07)	14 (-04)	-23 (-11)
OVERALL COURSE RATING	36 (33)	25 (29)	16 (09)	12 (08)	09 (02)	12 (16)	13 (-08)	14 (27)	08 (16)
II ENTHUSIASM									
Enthusiastic about Teaching	15 (29)	55 (42)	16 (00)	07 (02)	21 (15)	10 (00)	05 (16)	01 (09)	05 (06)
Dynamic & Energetic	08 (03)	70 (70)	15 (01)	11 (06)	08 (05)	06 (05)	07 (16)	01 (05)	06 (03)
Enhanced Presentations with Humor	10 (04)	66 (58)	-04 (06)	05 (01)	13 (02)	12 (02)	14 (07)	02 (-18)	-07 (-10)
Teaching Style Held Your Interest	09 (12)	59 (64)	23 (20)	16 (06)	06 (00)	03 (14)	10 (05)	06 (03)	-02 (-03)
OVERALL INSTRUCTOR RATING	12 (27)	40 (54)	23 (09)	14 (08)	23 (02)	11 (16)	10 (-08)	05 (27)	05 (16)
III ORGANIZATION									
Instructor Explanations Clear	12 (00)	07 (24)	55 (42)	20 (09)	05 (04)	10 (06)	13 (01)	06 (23)	-08 (-03)
Course Materials Prepared & Clear	06 (06)	03 (-02)	73 (69)	09 (01)	10 (-02)	09 (04)	06 (03)	10 (03)	01 (12)
Objectives Stated & Pursued	19 (12)	-05 (-08)	49 (41)	03 (05)	08 (05)	14 (08)	25 (27)	06 (05)	06 (06)
Lectures Facilitated Note Taking	-03 (02)	20 (09)	58 (53)	-17 (07)	-02 (05)	14 (04)	15 (06)	08 (01)	-04 (-05)
IV GROUP INTERACTION									
Encouraged Class Discussions	04 (06)	10 (02)	01 (03)	84 (86)	03 (00)	00 (00)	06 (00)	06 (-05)	00 (-03)
Students Shared Ideas/Knowledge	02 (08)	06 (-07)	-04 (-01)	85 (88)	05 (13)	05 (01)	08 (-02)	08 (-10)	-02 (01)
Encouraged Questions & Answers	03 (-04)	06 (09)	14 (06)	62 (69)	16 (-02)	15 (03)	07 (11)	08 (21)	00 (01)
Encouraged Expression of Ideas	07 (01)	02 (06)	01 (-11)	73 (75)	20 (09)	05 (07)	09 (12)	05 (09)	00 (-02)
V INDIVIDUAL RAPPORT									
Friendly Towards Students	-04 (10)	17 (06)	00 (-06)	13 (12)	68 (78)	-01 (-05)	13 (02)	10 (-05)	-07 (01)
Welcomed Seeking Help/Advice	04 (-10)	05 (02)	02 (07)	06 (00)	85 (75)	-04 (04)	12 (06)	05 (20)	03 (-04)
Interested in Individual Students	07 (10)	11 (09)	00 (01)	14 (07)	69 (77)	-01 (-09)	14 (03)	08 (-09)	03 (09)
Accessible to Individual Students	02 (-13)	-11 (-11)	16 (09)	09 (-02)	62 (43)	20 (25)	08 (13)	00 (14)	04 (07)
VI BREADTH OF COVERAGE									
Contrasted Implications	-05 (02)	12 (01)	05 (03)	08 (01)	-03 (01)	72 (84)	08 (-03)	14 (02)	08 (-06)
Gave Background of Ideas/Concepts	08 (03)	08 (10)	16 (07)	-03 (-02)	02 (-02)	71 (78)	01 (08)	11 (-01)	03 (03)
Gave Different Points of View	04 (-06)	04 (09)	11 (11)	08 (16)	06 (01)	72 (55)	07 (17)	01 (-06)	04 (08)
Discussed Current Developments	23 (29)	08 (-04)	-04 (-04)	05 (12)	09 (00)	50 (48)	06 (05)	16 (10)	-01 (-02)
VII EXAMINATIONS/GRADING									
Examination Feedback Valuable	-03 (01)	08 (09)	06 (-11)	09 (05)	08 (12)	-04 (03)	72 (62)	05 (-03)	09 (03)
Eval Methods Fair/Appropriate	06 (02)	00 (-03)	03 (14)	07 (06)	14 (00)	10 (17)	69 (64)	11 (11)	-08 (04)
Tested Emphasized Course Content	08 (00)	-01 (04)	11 (21)	01 (01)	06 (00)	11 (-04)	70 (58)	07 (10)	-02 (-03)
VIII ASSIGNMENTS									
Readings/Texts Valuable	-06 (09)	-03 (-03)	03 (07)	-01 (-06)	03 (01)	07 (-07)	01 (11)	91 (70)	02 (04)
Added to Course Understanding	12 (01)	-01 (-12)	01 (04)	09 (21)	01 (17)	-02 (08)	07 (05)	81 (56)	06 (10)
IX WORKLOAD/DIFFICULTY									
Course Difficulty (Easy-Hard)	-06 (00)	06 (-01)	04 (-05)	-04 (02)	-01 (00)	08 (00)	-04 (08)	10 (04)	85 (74)
Course Workload (Light-Heavy)	14 (-04)	-09 (-01)	03 (02)	07 (05)	00 (04)	06 (01)	00 (01)	00 (-04)	88 (86)
Course Pace (Too Slow - Too Fast)	-20 (07)	12 (00)	04 (18)	-12 (-09)	06 (02)	-03 (-07)	04 (08)	05 (-04)	62 (32)
Hours/week Outside of Class	14 (00)	07 (00)	-11 (00)	07 (02)	00 (02)	-04 (03)	03 (-08)	05 (21)	73 (46)

NOTE: Factor loadings in boxes are the loadings for items designed to measure each factor. All loadings are presented without decimal points. Factor analyses of student ratings and instructor self ratings (loadings in parentheses) consisted of a principal-components analysis, Kaiser normalization, and rotation to a direct oblimin criterion. The first nine unrotated factors for the instructor self ratings had eigenvalues of 9.5, 2.9, 2.5, 2.2, 2.0, 1.4, 1.3, 1.1 & 1.0, and accounted for 68% of the variance. For the student ratings the first nine eigenvalues were 19.9, 3.3, 2.3, 1.5, 1.2, 0.9, 0.7, 0.6 & 0.5, and accounted for 88% of the variance. The analyses were performed with the commercially available SPSS routine (See Nie, et. al., 1975).

APPENDIX IV

Multitrait-Multimethod Matrix: Correlations Between Student and Faculty Self Evaluations in All 329 Courses

INSTRUCTOR SELF EVALUATION FACTORS	INSTRUCTOR SELF-EVALUATION FACTORS								
	LEARN	ENTHU	ORGAN	GROUP	INDIV	BRDTH	EXAMS	ASIGN	WRKLD
LEARNING/VALUE	(83)								
ENTHUSIASM	29	(82)							
ORGANIZATION	12	01	(74)						
GROUP INTERACTION	01	03	-15	(90)					
INDIVIDUAL RAPPORT	-07	-01	07	02	(82)				
BREADTH	13	12	13	11	-01	(84)			
EXAMINATIONS	-01	08	26	09	15	20	(76)		
ASSIGNMENTS	24	-01	17	05	22	09	22	(70)	
WORKLD/DIFFICULTY	03	-01	12	-09	06	-04	09	21	(70)

STUDENT EVALUATION FACTORS	INSTRUCTOR SELF-EVALUATION FACTORS									STUDENT EVALUATION FACTORS								
	LEARN	ENTHU	ORGAN	GROUP	INDIV	BRDTH	EXAMS	ASIGN	WRKLD	LEARN	ENTHU	ORGAN	GROUP	INDIV	BRDTH	EXAMS	ASIGN	WRKLD
LEARNING/VALUE	(46)	10	-01	08	-12	09	-04	08	02	(95)								
ENTHUSIASM	21	(54)	-04	-01	-02	-01	-03	-09	-09	45	(96)							
ORGANIZATION	17	13	(30)	-03	04	07	09	00	-05	52	49	(93)						
GROUP INTERACTION	19	05	-20	(52)	00	-02	-14	-04	-08	37	30	21	(98)					
INDIVIDUAL RAPPORT	03	03	-05	13	(28)	-19	-03	-02	00	22	35	33	42	(96)				
BREADTH	26	15	09	00	-14	(42)	00	09	02	49	34	56	17	15	(94)			
EXAMINATIONS	18	09	01	-01	06	-09	(17)	-02	-06	48	42	57	34	50	33	(93)		
ASSIGNMENTS	20	03	02	09	-01	04	-01	(45)	12	52	21	34	30	29	40	42	(92)	
WORKLD/DIFFICULTY	-06	-03	04	00	03	-03	12	22	(69)	06	02	-05	-05	08	18	-02	20	(87)

NOTE: Values in the diagonals of the upper left and lower right matrices, the two triangular matrices, are reliability (coefficient alpha) coefficients (See Nie, et. al., 1977). Values in the diagonal of lower left matrix, the square matrix, are convergent validity coefficients that have been corrected for unreliability according to the Spearman Brown equation. The nine uncorrected validity coefficients, starting with Learning would be .41, .48, .25, .46, .25, .37, .13, .36, & .54. All correlation coefficients are presented without decimal point. Correlations greater than .10 are statistically significant

APPENDIX V

Correlation Matrix of Student and Instructor Evaluations of Teaching Effectiveness: Variables Include the Nine Evaluation Factor Scores, Overall Course Rating(OVCRS), Overall Instructor Rating (OVINS), and Students' Prior Subject Interest(PSI) As Rated By Instructors(INS) and Students(STD)

Instructor Self-Evalutions Correlated Insturctor & Student Ratings

	INS-LEARN	INS-ENTHU	INS-ORGAN	INS-GROUP	INS-INDIV	INS-BRDTH	INS-EXAMS	INS-ASIGN	INS-WORKD	INS-OVCRS	INS-OVINS	INS-PSI
INS-LEARN	1.00000	0.28629	0.11705	0.01442	-0.07000	0.12672	-0.00753	0.24350	0.02977	0.63415	0.39944	0.37003
INS-ENTHU	0.28629	1.00000	0.01016	0.03013	-0.01890	0.12388	0.08002	-0.00835	-0.01019	0.42961	0.63832	0.17626
INS-ORGAN	0.11705	0.01016	1.00000	-0.14690	0.07229	0.12855	0.26204	0.16684	0.11565	0.30036	0.21813	-0.07686
INS-GROUP	0.01442	0.03013	-0.14690	1.00000	0.02028	0.10656	0.08464	0.04569	-0.09180	0.10237	0.06652	0.02517
INS-INDIV	-0.07000	-0.01890	0.07229	0.02028	1.00000	-0.01411	0.14683	0.21794	0.05873	0.03166	0.10400	-0.03359
INS-BRDTH	0.12672	0.12388	0.12855	0.10656	-0.01411	1.00000	0.20264	0.08514	-0.04082	0.25064	0.30471	-0.05081
INS-EXAMS	-0.00753	0.08002	0.26204	0.08464	0.14683	0.20264	1.00000	0.21844	0.09085	0.25172	0.15981	-0.03206
INS-ASIGN	0.24350	-0.00835	0.16684	0.04569	0.21794	0.08514	0.21844	1.00000	0.21429	0.44191	0.19497	0.17523
INS-WORKD	0.02977	-0.01019	0.11565	-0.09180	0.05873	-0.04082	0.09085	0.21429	1.00000	0.23250	0.11887	0.15840
INS-OVCRS	0.63415	0.42961	0.30036	0.10237	0.03166	0.25064	0.25172	0.44191	0.23250	1.00000	0.57093	0.26106
INS-OVINS	0.39944	0.63832	0.21813	0.06652	0.10400	0.30471	0.15981	0.19497	0.11887	0.57093	1.00000	0.13211
INS--PST	0.37003	0.17626	-0.07686	0.02517	-0.03359	-0.05081	-0.03206	0.17523	0.15840	0.26106	0.13211	1.00000
STD-LEARN	0.40507	0.09896	-0.01235	0.08213	-0.12103	0.09272	-0.03511	0.07696	0.01615	0.22932	0.21306	0.29126
STD-ENTHU	0.21446	0.47635	-0.04084	-0.01306	-0.01647	-0.00739	-0.03248	-0.08968	-0.09797	0.17891	0.34079	0.10063
STD-ORGAN	0.17116	0.13154	0.25380	-0.03417	0.03600	0.06689	0.08563	-0.00296	-0.04776	0.14493	0.21633	0.04902
STD-GROUP	0.19201	0.04633	-0.20323	0.45420	-0.00386	-0.02418	-0.13744	-0.04718	-0.08077	0.09012	0.09382	0.14233
STD-INDIV	0.02909	0.03052	-0.05303	0.13084	0.24982	-0.18777	-0.03552	-0.02425	0.00253	0.06292	0.07843	0.08053
STD-BRDTH	0.25583	0.14879	0.08668	-0.00333	-0.14204	0.36743	0.00188	0.09469	0.02142	0.30286	0.31007	0.16181
STD-EXAMS	0.18295	0.08907	0.01684	-0.00682	0.05525	-0.08821	0.13476	-0.01878	-0.05586	0.08987	0.15953	0.01801
STD-ASIGN	0.20730	0.02612	0.01761	0.08588	-0.00579	0.04367	-0.01591	0.35554	0.12155	0.23583	0.24709	0.17414
STD-WORKD	-0.05769	-0.02929	0.04050	0.00068	0.03159	-0.03060	0.12315	0.22525	0.53949	0.15745	0.10205	0.06901
STD-OVCRS	0.36017	0.20144	0.02153	0.06766	-0.08998	0.04116	-0.01750	0.05658	0.02110	0.25793	0.30380	0.23250
STD-OVINS	0.26263	0.27045	0.02448	0.08090	-0.00101	0.03119	-0.00762	0.00113	-0.00470	0.22744	0.32483	0.16011
STD-PSI	0.25599	0.14733	-0.09491	0.13751	-0.09943	0.06383	-0.10372	0.06722	0.07456	0.20702	0.18066	0.36932

Student Evaluations Correlated with Instructor & Student Ratings

	STD-LEARN	STD-ENTHU	STD-ORGAN	STD-GROUP	STD-INDIV	STD-BRDTH	STD-EXAMS	STD-ASIGN	STD-WORKD	STD-OVCRS	STD-OVINS	STD-PSI
INS-LEARN	0.40507	0.21446	0.17116	0.19201	0.02909	0.25583	0.18295	0.20730	-0.05769	0.36017	0.26263	0.25599
INS-ENTHU	0.09896	0.47635	0.13154	0.04633	0.03052	0.14879	0.08907	0.02612	-0.02929	0.20144	0.27045	0.14733
INS-ORGAN	-0.01235	-0.04084	0.25380	-0.20323	-0.05303	0.08668	0.01684	0.01761	0.04050	0.02153	0.02448	-0.09491
INS-GROUP	0.08213	-0.01306	-0.03417	0.45420	0.13084	-0.00333	-0.00682	0.08588	0.00068	0.06766	0.08090	0.13751
INS-INDIV	-0.12103	-0.01647	0.03600	-0.00386	0.24982	-0.14204	0.05525	-0.00579	0.03159	-0.08998	-0.00101	-0.09943
INS-BRDTH	0.09272	-0.00739	0.06689	-0.02418	-0.18777	0.36743	-0.08821	0.04367	-0.03060	0.04116	0.03119	0.06383
INS-EXAMS	-0.03511	-0.03248	0.08563	-0.13744	-0.03552	0.00188	0.13476	-0.01591	0.12315	-0.01750	-0.00762	-0.10372
INS-ASIGN	0.07696	-0.08968	-0.00296	-0.04718	-0.02425	0.09469	-0.01878	0.35554	0.22525	0.05658	0.00113	0.06722
INS-WORKD	0.01615	-0.09797	-0.04776	-0.08077	0.00253	0.02142	-0.05586	0.12155	0.53949	0.02110	-0.00470	0.07456
INS-OVCRS	0.22932	0.17891	0.14493	0.09012	0.06292	0.30286	0.08987	0.23583	0.15745	0.25793	0.22744	0.20702
INS-OVINS	0.21306	0.34079	0.21633	0.09382	0.07843	0.31007	0.15953	0.24709	0.10205	0.30380	0.32483	0.18066
I NS-PST	0.29126	0.10063	0.04902	0.14233	0.08053	0.16181	0.01801	0.17414	0.06901	0.23250	0.16011	0.36932
STD-LEARN	1.00000	0.45461	0.52842	0.36887	0.22231	0.49368	0.48145	0.52065	0.05836	0.85258	0.68523	0.38166
STD-ENTHU	0.45461	1.00000	0.49654	0.30461	0.34951	0.33924	0.41881	0.21193	0.01665	0.65599	0.78133	0.13140
STD-ORGAN	0.52842	0.49654	1.00000	0.21480	0.33379	0.56187	0.57150	0.34476	-0.04621	0.68019	0.75951	0.06282
STD-GROUP	0.36887	0.30461	0.21480	1.00000	0.42045	0.16547	0.34117	0.30465	-0.05431	0.47447	0.47821	0.26553
STD-INDIV	0.22231	0.34951	0.33379	0.42045	1.00000	0.15576	0.50432	0.28827	0.07955	0.44436	0.57088	0.05496
STD-BRDTH	0.49368	0.33924	0.56187	0.16547	0.15576	1.00000	0.33407	0.40265	0.17754	0.57731	0.56075	0.21022
STD-EXAMS	0.48145	0.41881	0.57150	0.34117	0.50432	0.33407	1.00000	0.42279	-0.02296	0.67057	0.67719	0.05134
STD-ASIGN	0.52065	0.21193	0.34476	0.30465	0.28827	0.40265	0.42279	1.00000	0.20393	0.60727	0.49508	0.24686
STD-WORKD	0.05836	0.01665	-0.04621	-0.05431	0.07955	0.17754	-0.02296	0.20393	1.00000	0.13295	0.08986	0.18678
STD-OVCRS	0.85258	0.65599	0.68019	0.47447	0.44436	0.57731	0.67057	0.60727	0.13295	1.00000	0.89221	0.29761
STD-OVINS	0.68523	0.78133	0.75951	0.47821	0.57088	0.56075	0.67719	0.49508	0.08986	0.89221	1.00000	0.21206
STD-PSI	0.38166	0.13140	0.06282	0.26553	0.05496	0.21022	0.05134	0.24686	0.18678	0.29761	0.21206	1.00000

APPENDIX VI

Set of Sixteen Student, Course, and Instructor Characteristics Considered in Study II

LEVEL OF INTEREST IN THE SUBJECT PRIOR TO THIS COURSE (1-Very Low... 3-Medium...5-Very High)--Mean class average response was 3.4.

WORKLOAD/DIFFICULTY--An evaluation factor score representing four items; high values refer to courses that are more difficult, have a heavier workload, are faster paced, and require more hours outside of class.

OVERALL GPA (1-Below 2.5; 2-2.5 to 3.0; 3-3.0 to 3.4; 4-3.4 to 3.7; 5-Above 3.7)--Mean class average response was 3.3 (i.e., slightly higher than a B average).

ENROLLMENT--The number of students in the course. Mean enrollment was 34.5.

TEACHER RANK (1-Teaching Assistant, 2-Lecturer, 3-Instructor, 4-Assistant Professor, 5-Associate Professor, 6-Full Professor)--The average was 4.3. (Note: Teaching assistants were excluded from the analysis, and very few teachers were lecturers or instructors.)

PERCENT STUDENTS MAJORING IN SAME DIVISION AS THE COURSE (e.g., the percentage of social science students in a social science course)--The class average was 49%.

EXPECTED GRADE (1-F, 2-D, 3-C, 4-B, 5-A)--Mean class average response was 4.2, slightly higher than a B average.

REASON FOR TAKING THE COURSE (a subset of five variables):

1) Major Requirement (42%)
2) Major Elective (24%)
3) General Interest Only (16%)
4) General Education Requirement (12%)
5) Minor/Related Field (5%)

(Note: Students were asked to select one of the above responses, but were also given the option of leaving it blank; thus there is an implicit sixth "other" category.)

CLASS LEVEL (a subset of four variables):

1) Mean Year In School (1-Freshman, 2-Sophomore, 3-Junior, 4-Senior, 5-Graduate)--Mean class average response was 3.2.
2) Percentage of Freshmen and Sophomores--Mean class average response was 25%.
3) Percentage of Juniors and Seniors--Mean class average response was 65%.
4) Course Level (1-Lower Division, 2-Upper Division)--Mean class average was 1.7.

(Note: The fact that 70% of the classes were upper division classes and that the class average percentage of juniors and seniors was 65% is somewhat misleading. The lower division courses, though fewer in number, had substantially larger enrollments and were taken primarily by freshmen and sophomores.)

Footnotes

Grateful acknowledgements are extended to John Schutz, Joseph Kertes, and Robert Linnell, of the University of Southern California, and Raymond Orbach, James Trent, and Robert Pace of the University of California, Los Angeles, who have supported the student evaluation program. Thanks are also extended to J. U. Overall, and Zoe Cosgrove, who made many helpful suggestions on earlier drafts of this paper. This study was also the basis of a paper presented at the 1980 Annual Meeting of the American Educational Research Association (Boston; April, 1980).

REFERENCES

Blalock, H. W. (ed.) Causal Models in the Social Sciences. Chicago: Aldine Atherton, 1971.

Brown, D. L. Faculty ratings and student grades: A university-wide multiple regression analysis. Journal of Educational Psychology, 1976, 68, 573-578.

Burton, D. Student ratings--an information source for decision-making. In Cope, R. G. (Ed.) Information for Decisions in Postsecondary Education: Proceedings of the 15th Annual Forum, Association for Institutional Research, St. Louis, Missouri, April, 1975. (Summary)

Centra, J. A. The Student Instructional Report: Item Reliabilities (Student Instructional Report No. 3) Princeton, N. J.: Educational Testing Service, 1973.

Centra, J. A. Student ratings of instruction and their relationship to student learning. American Educational Research Journal, 1977, 14, 17-24.

Cohen, J. & Cohen, P. Applied Multivariate Regression/Correlation Analysis for the Behavioral Sciences. New York: Wiley, 1975.

Cohen, S. A. & Berger, W. G. Dimensions of students' ratings of college instructors underlying subsequent achievement on course examinations. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 605-606. (Summary)

Cooley, W. W. & Lohnes, P. R. Multivariate Data Analysis. New York: Wiley, 1971.

Cooley, W. W. & Lohnes, P. R. Evaluation Research in Education. New York: Wiley, 1976.

Costin, F., Greenough, W. T. & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Dixon, W. J. (Ed.) Biomedical Computer Programs. Los Angeles: University of California Press, 1975.

Doyle, K. O., & Crichton, L. I. Student, peer, and self evaluations of college instructors. Journal of Educational Psychology, 1978, 70, 815-826.

Feldman, K. A. Grades and college students' evaluations of their courses and teachers. Research in Higher Education, 1976, 4, 69-111.

Feldman, K. A. Consistency and variability among college students in rating their teachers and courses: A review and analysis. Research in Higher Education, 1977, 6, 223-274.

Frey, P. W. Student ratings of teaching: Validity of several rating factors. Science, 1973, 182, 83-85.

Frey, P. W. A two-dimensional analysis of student ratings of instruction. Research in Higher Education, 1978, 9, 69-91.

Hildebrand, M., Wilson, R. C., & Dienst, E. R. Evaluating University Teaching. Berkeley: Center for Research and Development in Higher Education, University of California, Berkeley, 1971.

Howard, G.S. & Schmeck, R.R. Relationship of changes in student motivation to student evaluations of instruction. Research in Higher Education, 1979, 10, 305-315.

Kerlinger, F. N., & Pedhazur, E. J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart and Winston, 1973.

Marsh, H. W. The validity of students' evaluations: Classroom evaluations of instructors independently nominated as best and worst teachers by graduating seniors. American Educational Research Journal, 1977, 14, 441-447.

Marsh, H.W. The influence of student, course, and instructor characteristics on evaluations of university teaching. American Educational Research Journal, in press.

Marsh, H. W., Fleiner, H., & Thomas, C. S. Validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology, 1975, 67, 833-839.

Marsh, H. W., & Overall, J. U. Students' Evaluations of Instructional Effectiveness: Relationship to Student, Course, and Instructor Characteristics. Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, March 1978. (ERIC Document Reproduction Service No. ED 155 217)

Marsh, H. W. & Overall, J. U. Long-term stability of students' evaluations: A note on Feldman's "Consistency and variability among college students in rating their teachers and courses". Research in Higher Education, 1979a, 10, 139-147.

Marsh, H. W. & Overall, J. U. Validity of Students' Evaluations of Teaching: A Comparison With Instructor Self Evaluations By Teaching Assistants, Undergraduate Faculty and Graduate Faculty. Paper presented at Annual Meeting of the American Educational Research Association, April 1979b.

Marsh, H.W. & Overall, J.U. Validity of students' evaluations of teaching effectiveness: Cognitive and affective criteria. Journal of Educational Psychology, in press.

Marsh, H. W., Overall, J. U. & Kesler, S. P. Validity of student evaluations of instructional effectiveness: A comparison of faculty self-evaluations and evaluations by their student. Journal of Educational Psychology, 1979a, 71, 149-160.

Marsh, H. W., Overall, J. U. & Kesler, S. P. Class size, students' evaluations, and instructional effectiveness. American Educational Research Journal, 1979b, 16, 57-70.

Marsh, H. W., Overall, J. U., & Thomas, C. S. The Relationship Between Students' Evaluation of Instruction and Expected Grade. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, April 1976. (ERIC Document Reproduction Service No. ED 126140)

McKeachie, W. J. Correlates of student ratings. In Sockloff, A. L. (Ed.), Proceedings: The First Invitational Conference on Faculty Effectiveness as Evaluated by Students. Philadelphia: Measurement and Research Center, Temple University, 1973.

Menges, R. J. The new reporters: students rate instruction. In C. R. Pace (Ed.) Evaluating Learning & Teaching. San Francisco: Jossey-Bass, 1973.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D. H. Update to Statistical Package For the Social Sciences. New York: McGraw-Hill, 1977.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D. H. Statistical Package For the Social Sciences. New York: McGraw-Hill, 1975.

Overall, J. U., & Marsh, H. W. Cognitive and Affective Outcomes: Their relationship to effective teaching and students' evaluations. Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, March 1978.

Pohlmann, J. T. A multivariate analysis of selected class characteristics and student ratings of instruction. Multivariate Behavioral Research, 1975, 10, 81-92.

Price, J. A. & Magoon, A. J. Predictors of college student ratings of instructors. Proceedings of the 79th Annual Convention of the American Psychological Association, 1971, 6, 523-524.

Remmers, H. H. Teaching methods in research on teaching. In Gage (Ed.) Handbook on Teaching. Chicago: Rand McNally, 1963.

Webb, W. B., & Nolan, C. Y. Student, supervisor, and self-ratings of instructoral proficiency. Journal of Educational Psychology, 1955, 46, 42-46.